



INTERPOL's 18th International Forensic Science Managers Symposium



INTERPOL

12 October 2016

“State-of-the-Art” Forensic DNA

John M. Butler, PhD

National Institute of Standards and Technology

Gaithersburg, Maryland

United States of America



National Institute of Standards and Technology

- Science agency **part of the U.S. Department of Commerce**
- Started in 1901 as the **National Bureau of Standards**
- Name changed in 1988 to the **National Institute of Standards and Technology (NIST)**
- Forensic science research activities dating back to 1920s
- Partnership since 2013 with U.S. Department of Justice to create the National Commission on Forensic Science (NCFS) and the Organization of Scientific Area Committees (OSAC)

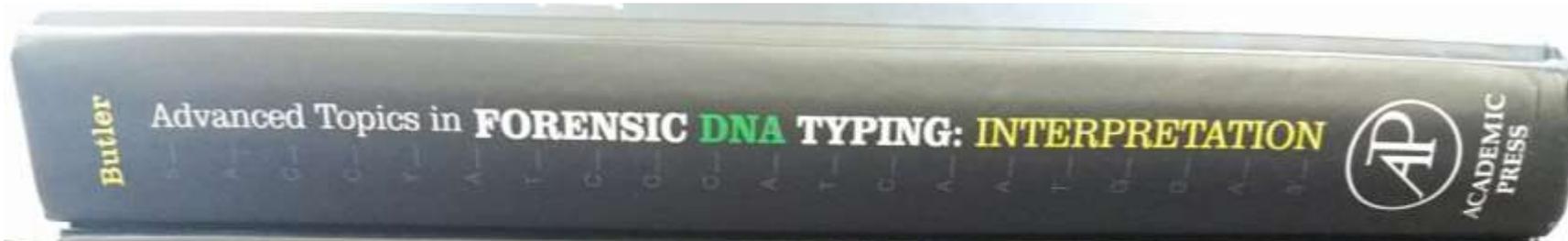
- Primary campus in Gaithersburg, Maryland (near Washington, D.C.)
- >3,400 employees and >3,700 associates
- Supplies >1300 reference materials
- Defines official time for the U.S.



DNA reference material

Butler Books on Forensic DNA Typing

2015



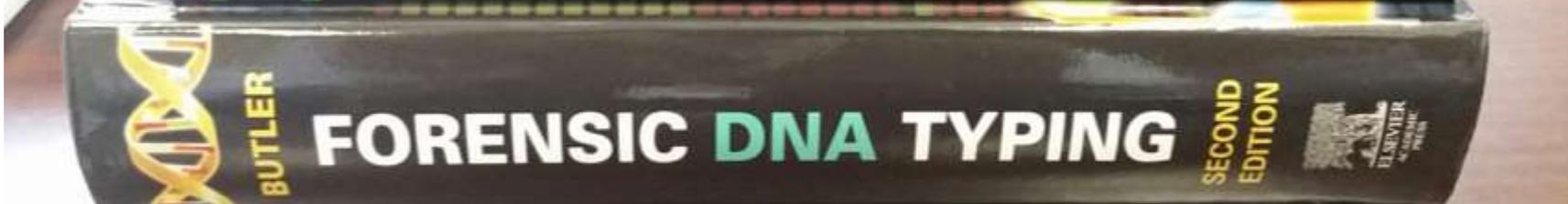
2012



2010



2005



2001



DNA Capabilities to Aid Forensic Investigations

1. The **ability to identify the perpetrator**
2. Weight-of-evidence based on established genetic principles and statistics (Hardy-Weinberg 1908)
3. Established characteristics of genetic inheritance enables close **biological relatives** to be used for reference points using kinship associations
4. Superb **sensitivity** with PCR amplification (opens the possibility for contamination)
5. Well-established **quality assurance measures**
6. New **technology development** aided by genomics

Successful interpretation of DNA (Q-to-K comparison) depends on quality of the crime scene evidence (Q) and availability of suitable reference samples (K)

Thoughts on the Future of Forensic DNA Published in 2015

PHILOSOPHICAL
TRANSACTIONS B

rstb.royalsocietypublishing.org



CrossMark
click for updates

Opinion piece

Cite this article: Butler JM. 2015 The future of forensic DNA analysis. *Phil. Trans. R. Soc. B* **370**: 20140252.

<http://dx.doi.org/10.1098/rstb.2014.0252>

Accepted: 26 February 2015

One contribution of 15 to a discussion meeting issue 'The paradigm shift for UK forensic science'.

The future of forensic DNA analysis

John M. Butler

National Institute of Standards and Technology, Gaithersburg, MD, USA

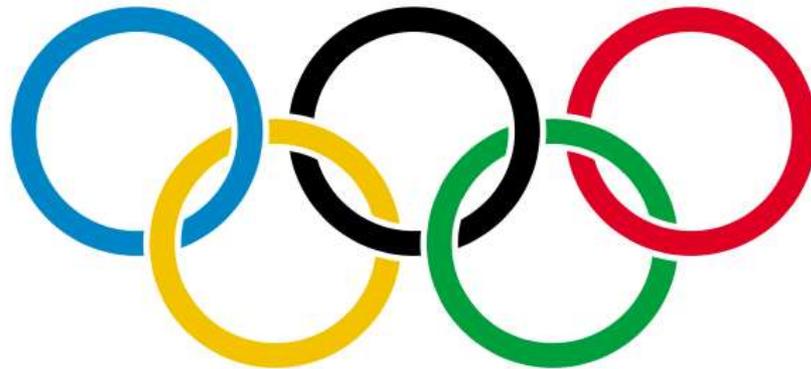
The author's thoughts and opinions on where the field of forensic DNA testing is headed for the next decade are provided in the context of where the field has come over the past 30 years. Similar to the Olympic motto of 'faster, higher, stronger', forensic DNA protocols can be expected to become more rapid and sensitive and provide stronger investigative potential. New short tandem repeat (STR) loci have expanded the core set of genetic markers used for human identification in Europe and the USA. Rapid DNA testing is on the verge of enabling new applications. Next-generation sequencing has the potential to provide greater depth of coverage for information on STR alleles. Familial DNA searching has expanded capabilities of DNA databases in parts of the world where it is allowed. Challenges and opportunities that will impact the future of forensic DNA are explored including the need for education and training to improve interpretation of complex DNA profiles.

**Addressed Rapid DNA and
Next-Generation Sequencing**

Current Trends in Forensic DNA

are Similar to the Olympic Motto of
Citius, Altius, Fortius

“Faster, Higher, Stronger”



Current Trends in Forensic DNA

- ***Faster results:*** Rapid DNA capabilities and new sample-to-answer integrated instruments
- ***Higher information content:*** Next-generation sequencing (NGS) for more markers & STR allele information
- ***Higher sensitivity:*** New assays lowering the limits of detection, which makes interpretation more challenging
- ***Stronger conclusions:*** Mixture interpretation with probabilistic genotyping models

Forensic Science International: Genetics

September 2015 Issue (Volume 18)



- Guest Editor:
John M. Butler
(NIST)
- 13 review articles
on New Trends in
Forensic Genetics
- Authors are from Austria,
Australia, Denmark, the
Netherlands, Norway,
Spain, the United Kingdom,
and the United States



From 2015 Special Issue: New Trends in Forensic Genetics

Forensic Science International: Genetics 18 (2015) 90–99



ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Review

Rapid PCR of STR markers: Applications to human identification

Erica L. Romsos^{*}, Peter M. Vallone

National Institute of Standards and Technology, 100 Bureau Drive, MS 8314, Gaithersburg, MD 20899-8314, USA

Cites 118 articles on Rapid DNA



From 2015 Special Issue: New Trends in Forensic Genetics

Forensic Science International: Genetics 18 (2015) 78–89



ELSEVIER

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Next generation sequencing and its applications in forensic genetics

Claus Børsting*, Niels Morling

Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

Cites 94 articles on Next Generation Sequencing

Acknowledgment and Disclaimers

Research at NIST on Rapid DNA and Next-Generation Sequencing is a partnership with the FBI Laboratory conducted within the **Applied Genetics Group** led by Peter Vallone, Katherine Gettings, and Erica Romsos with funding in part through the FBI Biometrics Center of Excellence

I have been fortunate to have had discussions with numerous scientists on interpretation issues including Mike Coble, Bruce Heidebrecht, Robin Cotton, Charlotte Word, Catherine Grgicak, Peter Gill, Ian Evett, John Buckleton, Hari Iyer, Steve Lund ...

Points of view are mine and do not necessarily represent the official position or policies of the US Department of Justice or the National Institute of Standards and Technology.

Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. **In no case does such identification imply a recommendation or endorsement** by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments or equipment identified are necessarily the best available for the purpose.

Current Forensic DNA Testing

- **Short tandem repeat (STR) markers** are used
 - Typically **15 to 22 STRs examined** with commercial kits (e.g., Identifiler, PowerPlex 16, NGM, GlobalFiler, Fusion)
- STR length (and sequence) varies among individuals
 - DNA molecules are **labeled with fluorescent dyes and separated by size using CE** (capillary electrophoresis)
 - **Only the STR length is measured** against an internal size standard and calibrated with an allelic ladder (which is a combination of the most common possibilities of alleles)
- **National DNA databases** using STR markers now exist in >50 countries (>75 million STR profiles total)
 - **Having core STR markers in common is critical** to enable comparisons across laboratories and between countries

DNA profile with 15 STR markers and sex-typing (Identifiler STR kit)

1-2 days

Steps Involved

Collection

Characterization

Extraction

Quantitation

Amplification

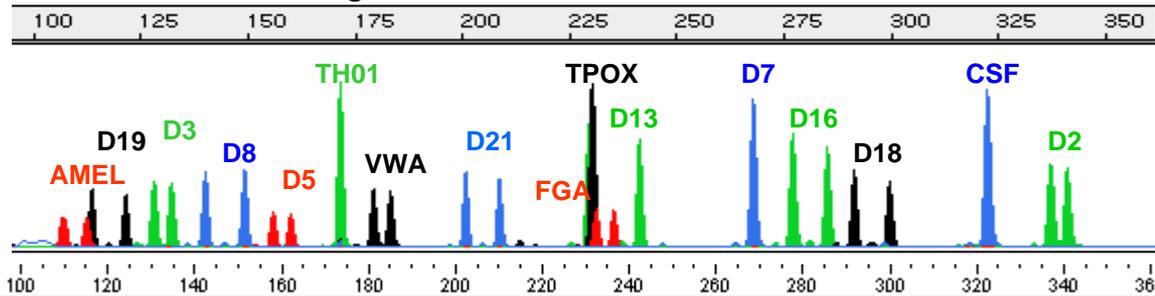
CE Analysis

Interpretation

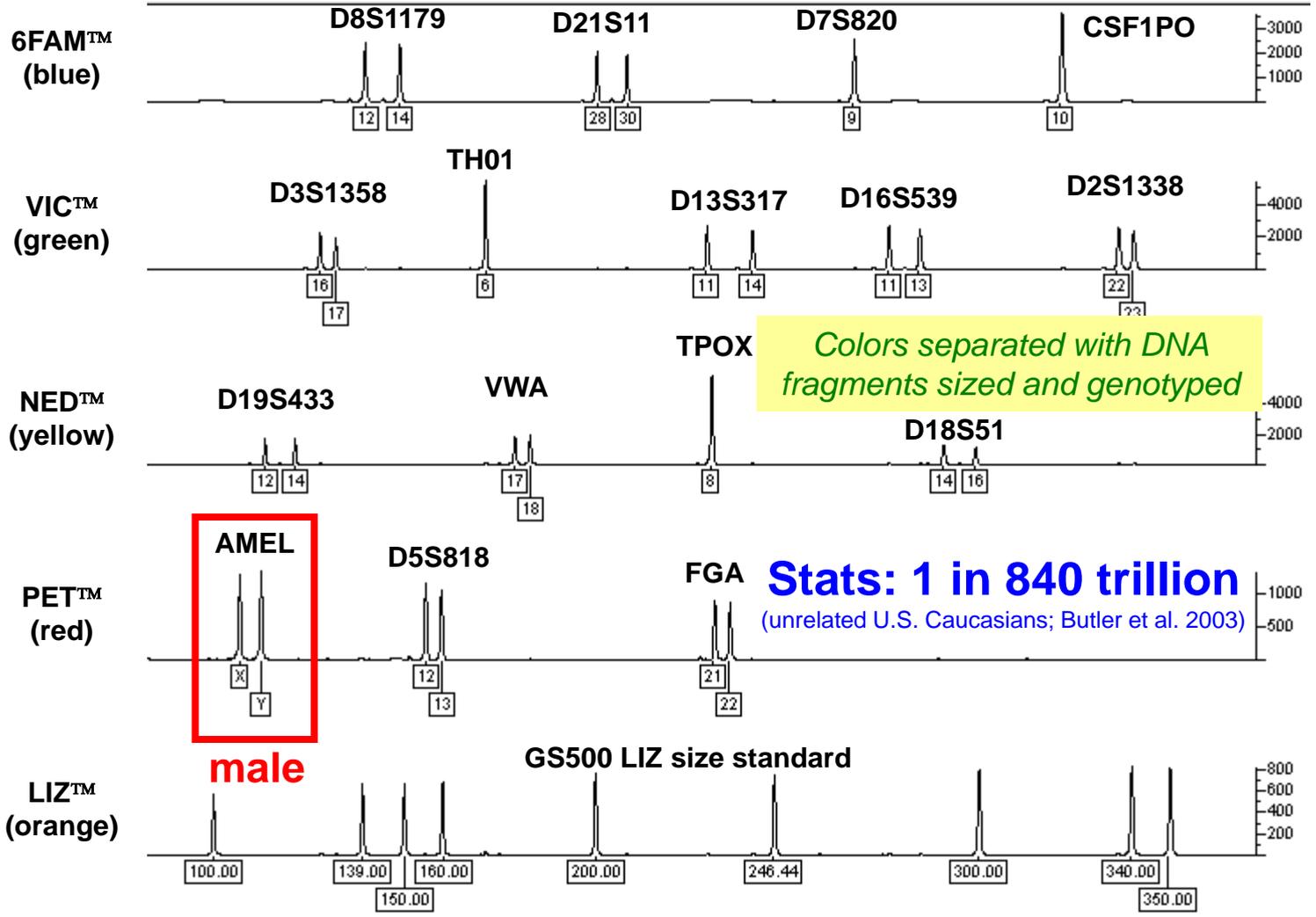
Statistics

Report

Size range: 100 nucleotides to 350 nucleotides



Colors overlaid



Colors separated with DNA fragments sized and genotyped

Stats: 1 in 840 trillion
(unrelated U.S. Caucasians; Butler et al. 2003)

male

Rapid DNA

- Faster results opens up potential new applications
 - DNA testing at embassies, border crossings, or police booking stations
- Two commercial sources:
 - RapidHIT (IntegenX)
 - DNAScan (NetBio/GE Health)
- NIST studies and published validation work

Early Demonstration of Rapid DNA

NIST
research
published in
December
2008

Forensic Science International: Genetics 3 (2008) 42–45

Contents lists available at ScienceDirect

 Forensic Science International: Genetics 

journal homepage: www.elsevier.com/locate/fsig

Short communication

Demonstration of rapid multiplex PCR amplification involving 16 genetic loci[☆]

Peter M. Vallone^{*}, Carolyn R. Hill, John M. Butler

National Institute of Standards and Technology, Biochemical Science Division, 100 Bureau Drive, Mail Stop 8311, Gaithersburg, MD 20899-8311, United States

Table 1

Comparison of thermal cycling times.

| Parameter | Standard | Rapid |
|-----------------|----------|---------|
| Hot start (min) | 10 | 1 |
| Denature (s) | 60 | 5 |
| Anneal (s) | 60 | 10 |
| Elongate (s) | 60 | 10 |
| Soak (min) | 60 | 1 |
| Ramp rate (°/s) | 1 | 4 |
| Cycles | 28 | 28 |
| Time | 2:58:41 | 0:35:38 |

Forensic DNA testing involves copying segments of DNA with the polymerase chain reaction (PCR). Innovations in this work involved **use of new DNA polymerases** and **faster thermal cycling** with shorter dwell times for each step.

35 minutes instead of ~3 hours

2014

Rapid DNA Instruments

ANDE (NetBio)



- One biochipset
 - Stored at RT
 - Shelf life \approx 6 months
- RFID swabs tagged for sample tracking

PowerPlex 16 loci
 \approx 86 min runtime
(5 samples)

ANDE PP16

RapidHIT 200 (IntegenX)



- Kit = 4 components
 - Stored between RT-4°C
 - Shelf life \approx 6 months @ 4°C
- Cotton Swabs

PowerPlex 16 loci
 \approx 90 min runtime
(5 samples)

RH200 PP16

GlobalFiler Express loci
 \approx 120 min runtime
(1-7 samples)

RH200 GFE

Rapid DNA Maturity Assessment

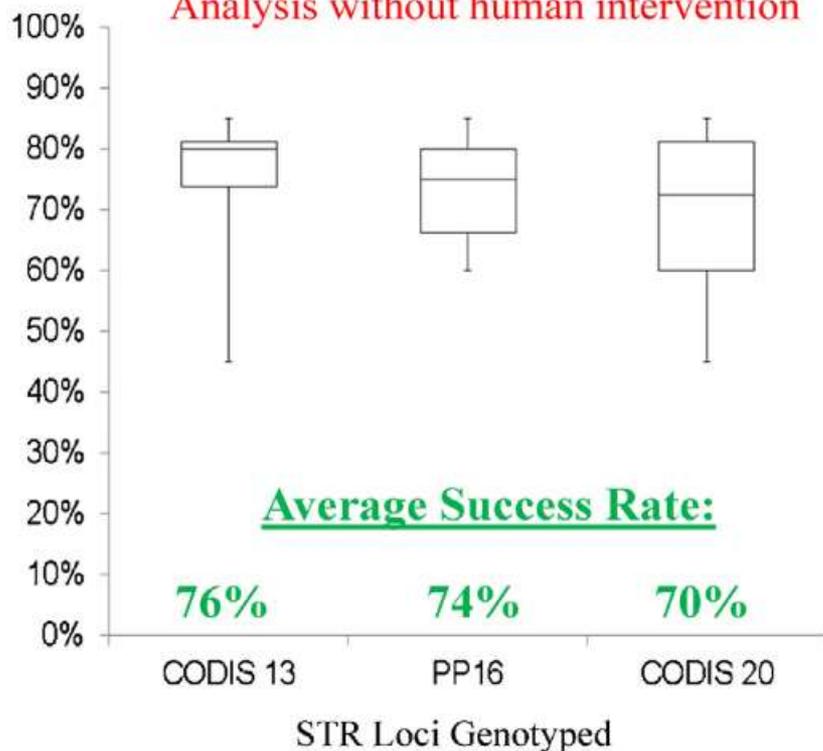
- 2014 Rapid DNA Maturity Assessment
 - (Poster) <https://www.nist.gov/sites/default/files/documents/mml/bmd/genetics/Romsos-ISFG-2015-Rapid-MA.pdf>
 - (Talk) http://www.cstl.nist.gov/strbase/pub_pres/Romsos_2014-Rapid-DNA-MA-Results-GIS2015.pdf
 - (Paper) [http://www.fsigeneticssup.com/article/S1875-1768\(15\)30166-9/pdf](http://www.fsigeneticssup.com/article/S1875-1768(15)30166-9/pdf)
- For more information regarding FBI-funded **NIST research with rapid DNA**, see <https://www.nist.gov/programs-projects/dna-biometrics>

| Rapid DNA Instrument Platforms | Number of Participating Labs | Total Instruments | Samples Attempted | Core CODIS Success (Rapid DNA Analysis) | Core CODIS Success (Modified Rapid DNA Analysis) |
|--------------------------------|------------------------------|-------------------|-------------------|---|--|
| 2 | 7 | 11 | 280 | 76.1% | 80.0% |

E.L. Romsos et al. / Forensic Science International: Genetics Supplement Series 5 (2015) e1–e2

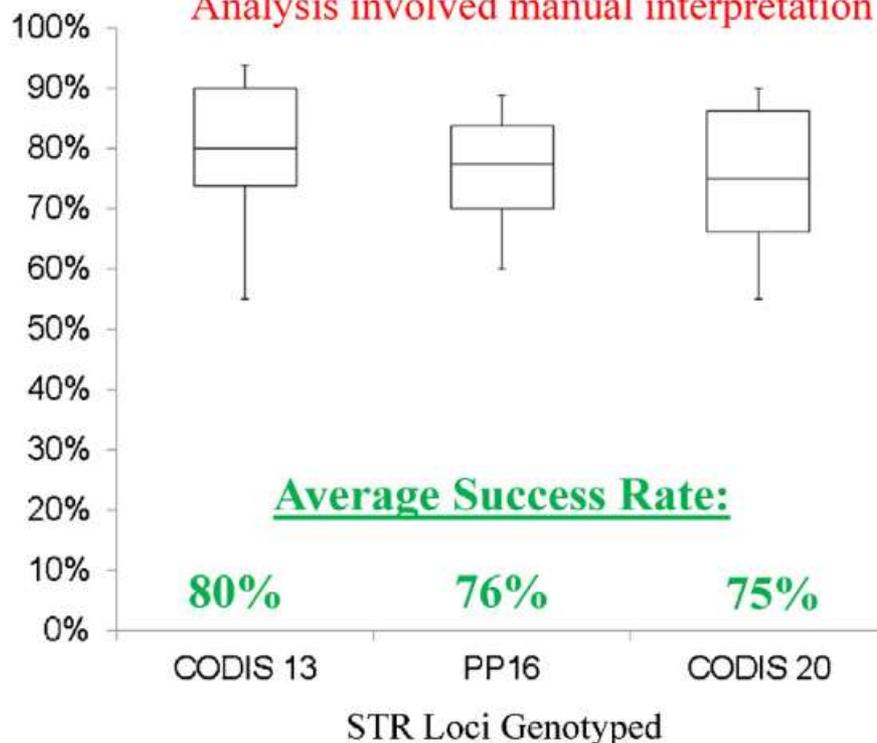
(a) Rapid DNA Analysis

Analysis without human intervention



(b) Modified Rapid DNA Analysis

Analysis involved manual interpretation



2 labs, 250 samples, 88% success rate

Forensic Science International: Genetics (May 2015) 16:181-194

Developmental validation of a fully integrated sample-to-profile rapid human identification system for processing single-source reference buccal samples

Stevan Jovanovich ^{a,*}, Greg Bogdan ^a, Richard Belcinski ^a, Jacklyn Buscaino ^a, Dean Burgi ^a, Erica L.R. Butts ^b, Kaiwan Chear ^a, Brian Ciopyk ^a, David Eberhart ^a, Omar El-Sissi ^a, Helen Franklin ^a, Stefanie Gangano ^a, Jennifer Gass ^a, Dennis Harris ^a, Lori Hennessy ^a, Alex Kindwall ^a, David King ^a, Jim Klevenberg ^a, Yuan Li ^a, Neelima Mehendale ^a, Roger McIntosh ^a, Bill Nielsen ^a, Charles Park ^a, Francesca Pearson ^a, Robert Schueren ^a, Nancy Stainton ^a, Charles Troup ^a, Peter M. Vallone ^b, Mattias Vangbo ^a, Timothy Woudenberg ^a, David Wyrick ^a, Stephen Williams ^a

^a IntegenX Inc. 5720 Stoneridge Drive, Suite 300, Pleasanton, CA 94588-2739, USA

^b National Institute of Standards and Technology, Gaithersburg, MD 20899-8314, USA

RapidHIT 200 System
(up to 5 samples at a time)





<http://www.integenx.com>

RapidHIT Systems



RapidHIT 200 System
(up to 7 samples at a time now)



Image courtesy of IntegenX (Kevin Bekak)

RapidHIT ID System
(1 sample at a time)

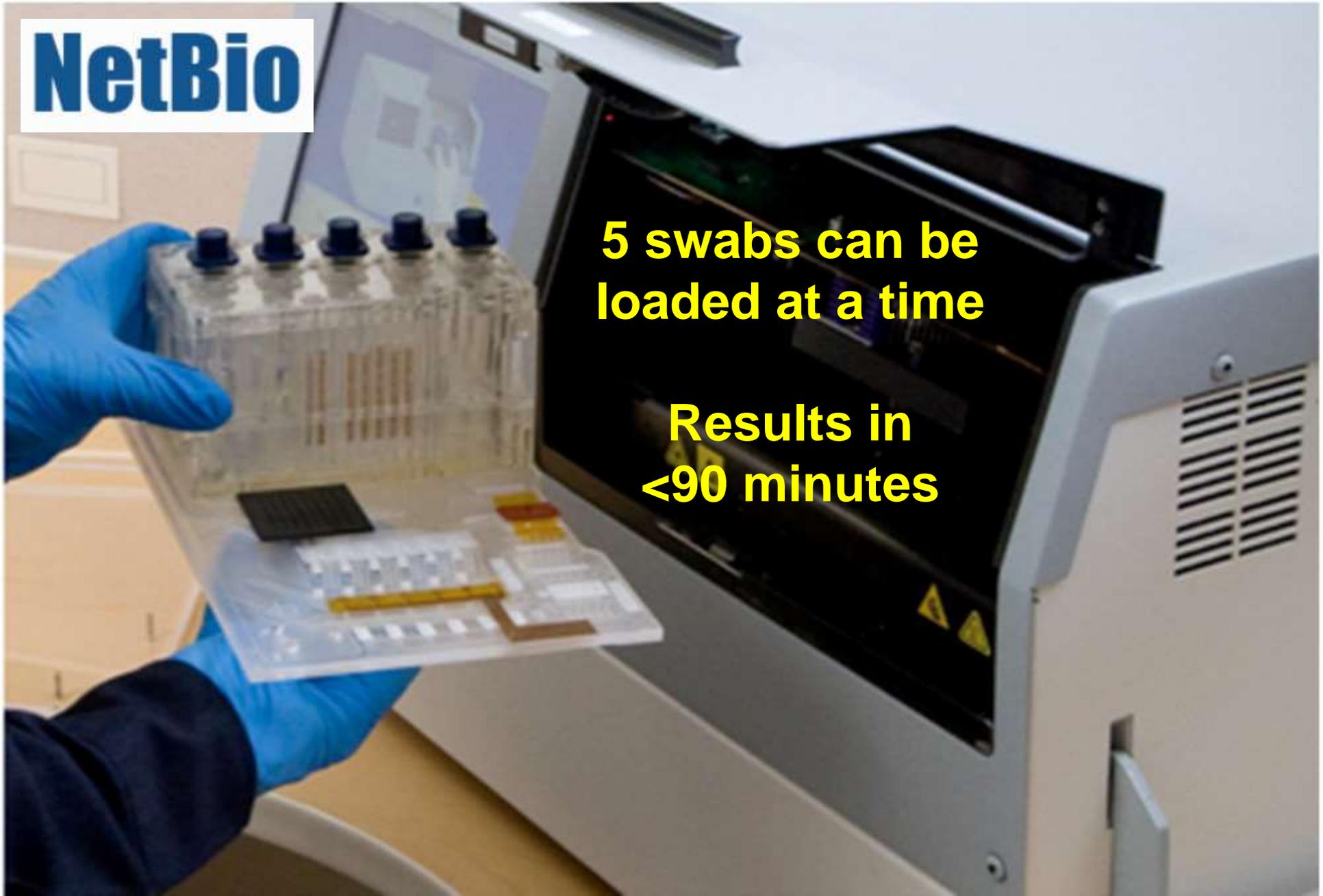
Multiple Units of the RapidHIT ID

Image courtesy of IntegenX (Kevin Bebak)





NetBio



**5 swabs can be
loaded at a time**

**Results in
<90 minutes**

Fig. 1. The BioChipSet Cassette can be loaded into the DNAScan instrument by non-technical personnel.

8 labs, 1362 samples, >2300 swabs examined
99.9% accuracy, 84% success rate (91% with human review)

Forensic Science International: Genetics (September 2016) 25:145-156

Research paper

Developmental validation of the DNAscan™ Rapid DNA Analysis™
instrument and expert system for reference sample processing

Angelo Della Manna^a, Jeffrey V. Nye^b, Christopher Carney^c, Jennifer S. Hammons^d,
Michael Mann^d, Farida Al Shamali, PhD^e, Peter M. Vallone, PhD^f, Erica L. Romsos, PhD^f,
Beth Ann Marne^g, Eugene Tan, PhD^h, Rosemary S. Turingan, PhD^h, Catherine Hogan^h,
Richard F. Selden, MD PhD^h, Julie L. French^{i,*}

^a Alabama Department of Forensic Sciences, 2026 Valleydale Road, Hoover, AL 35244, USA

^b Michigan State Police, 7320 North Canal Road, Lansing, MI 48913, USA

^c Florida Department of Law Enforcement DNA Investigative Support Database, 2331 Phillips Road, Tallahassee, FL 32308, USA

^d Defense Forensic Science Center, Office of Chief Scientist, 4930 North 31st Street, Forest Park, GA 30297, USA

^e Dubai Police GHQ, Gen. Dept. Forensic Sciences & Criminology, P.O. Box 1493, Dubai, UAE

^f National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8314, USA

^g Pennsylvania State Police, Forensic DNA Division, 80N. Westmoreland Avenue, Greensburg, PA 15601, USA

^h NetBio, 830 Winter Street, Waltham, MA, USA¹

ⁱ GE Healthcare Life Sciences, 100 Results Way, Marlborough, MA 01752, USA

DNAscan System
(up to 5 samples at a time)

NetBio

Summary of Rapid DNA

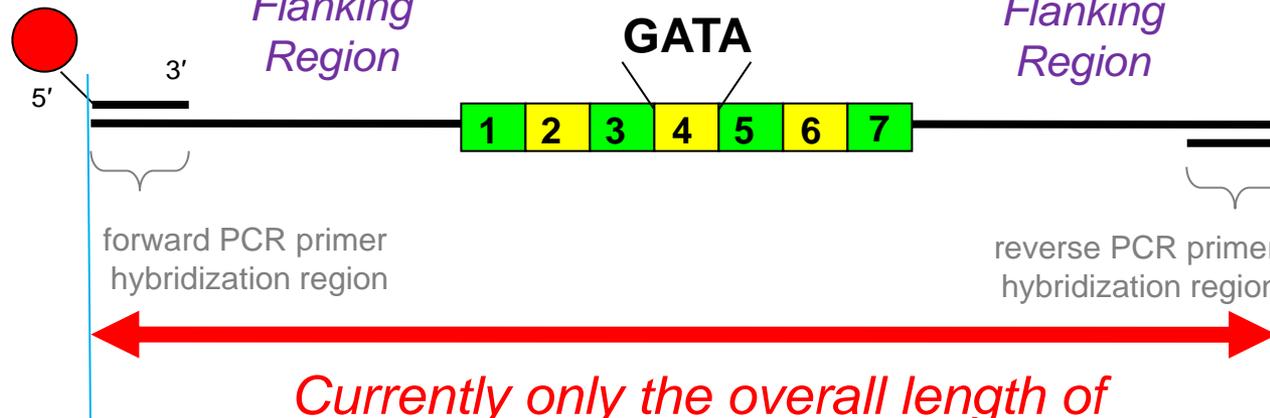
- Integrated instruments (sample-to-result) can produce reliable DNA results in <90 minutes
- Size-based analysis of 15 to 22 STR markers
- Success rates were typically >80%
- Reagent costs are approximately 10 times conventional testing (≈\$300 per sample)
 - But do not have to maintain trained analysts or full DNA laboratory to generate results

Next-Generation Sequencing (NGS) or Massively Parallel Sequencing (MPS)

- Higher information content opens up potential new applications
 - DNA testing with single nucleotide polymorphisms (SNPs) and more STRs, biogeographical ancestry, phenotyping, and possible improved mixture resolution (from ability to see STR allele sequence differences)
- Two commercial sources:
 - MiSeq FGx (Illumina)
 - Ion PGM or Ion S5 (ThermoFisher Scientific)
- NIST studies and published validation work

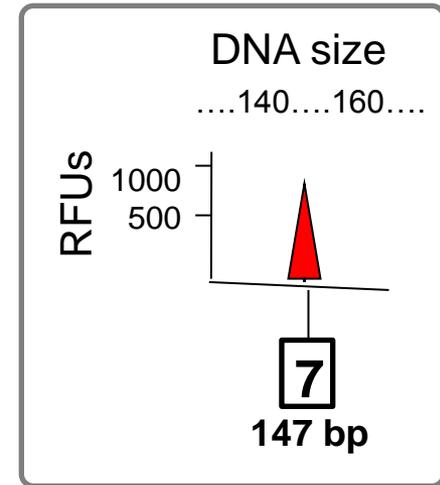
Short Tandem Repeat (STR) Analysis

Fluorescent dye-labeled primer



Currently only the overall length of the STR DNA fragment is measured

CE Result



Full DNA sequence analysis enables observation of potential differences in the flanking regions and the STR repeat

Forward primer TCCAAGCTCTTCC

Flanking Region TCTTCCCTAGAT[C/T]AATACAGACAGAAGACAGGTG

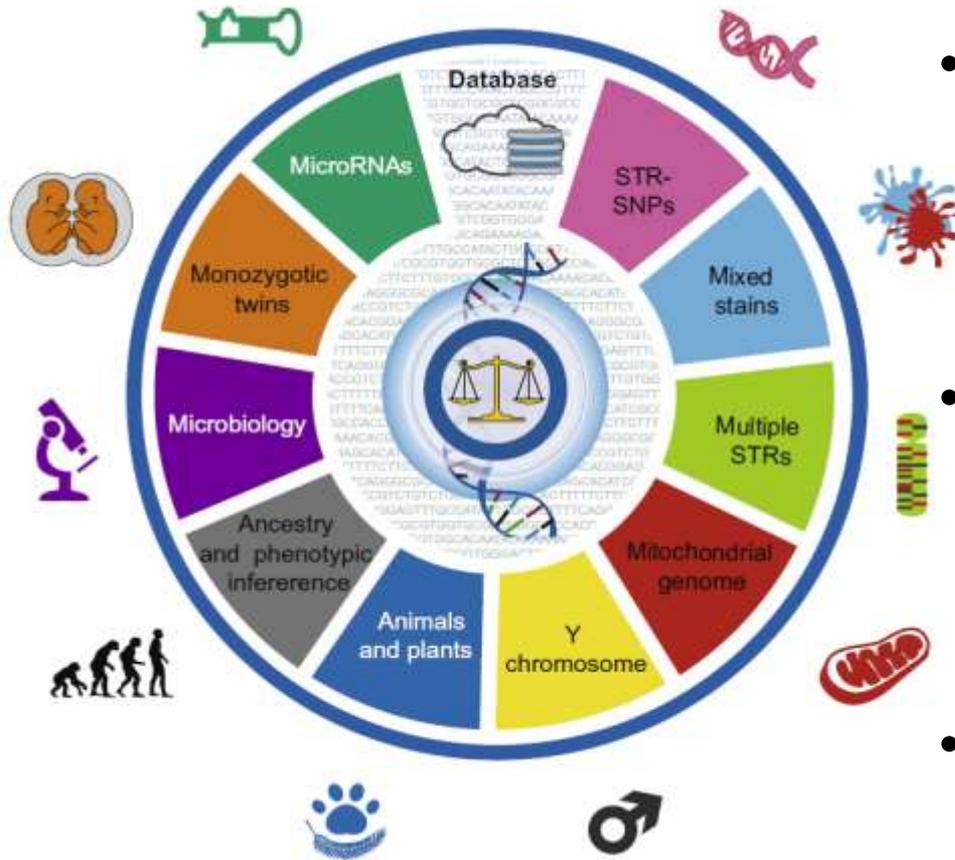
STR Repeat **GATA GATA GATA GATA GA[T/C]A GATA GATA**

Flanking Region TCATTGA[A/G]AGACAAAACAGAGATGGAT[G/A]ATA

Reverse primer TACAGATGCACAC

Forensic Use of NGS/MPS

- **More information content from STR allele sequences**
- **More markers can be simultaneously tested** along with autosomal STRs (e.g., SNPs, Y-STRs, X-STRs, mtDNA)
- **Additional applications are possible** (e.g., ancestry and phenotyping inference possible with SNPs)
- **New capabilities such as resolution of twins** with full genome sequencing





MiSeq FGx Forensic Genomics System



MiSeq FGx

ForenSeq DNA Signature Prep Kit

231 Markers Examined
(58 STRs + 173 SNPs)

Amplicon sizes

27 Autosomal STRs

61 to 467 bp

24 Y-chromosome STRs

119 to 390 bp

7 X-chromosome STRs

157 to 462 bp

95 Identity SNPs

63 to 231 bp

22 Phenotyping SNPs

73 to 227 bp

56 Ancestry SNPs

67 to 200 bp

*Many markers can be run simultaneously
Short amplicons enables better results with degraded DNA*

<http://www.illumina.com/systems/miseq-fgx.html>

Precision ID NGS System for Human Identification



Ion PGM

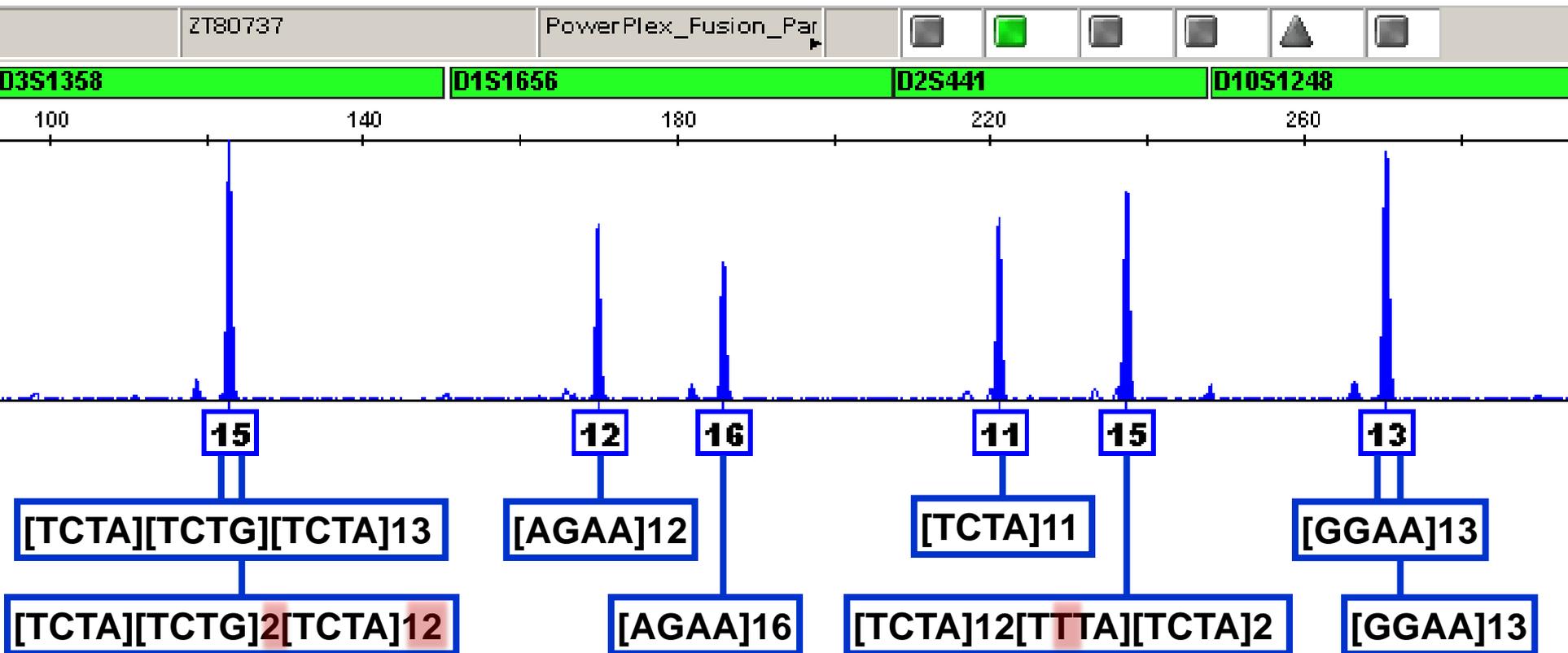


Ion S5

| Precision ID Panel | Markers Examined | Amplicon sizes (average length) |
|----------------------|--|---------------------------------------|
| GlobalFiler NGS | 30 autosomal STRs, one Y indel, Amelogenin X & Y | 129 to 250 bp |
| Ancestry | 165 SNPs | 120 to 130 bp |
| Identity | 113 SNPs | 132 to 141 bp |
| mtDNA Whole Genome | 16,569 bp mtGenome | 163 bp with amplicon overlap of 11 bp |
| mtDNA Control Region | 1.2 kb (16024 to 574) | 153 bp with amplicon overlap of 18 bp |

*Many markers can be run simultaneously
Short amplicons enables better results with degraded DNA*

Forensic STR Sequence Diversity



Sequence-Based Heterozygote: A locus that appears homozygous in length-based measurements (such as CE), but is heterozygous by sequence

Sequence Variability in STR Alleles Across 183 Samples

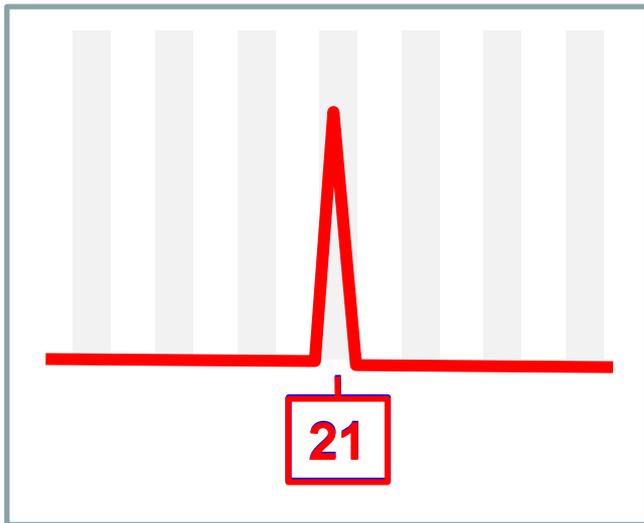
Sequence data provides further information with these 6 STR loci

No additional information with sequence data with these 6 STR loci

| STR Locus | CE (length only) | NGS (with sequence) | STR Locus | CE (length only) | NGS (with sequence) | |
|-----------|------------------|---------------------|-----------|------------------|---------------------|----|
| D12S391 | 17 | +36 | 53 | D22S1045 | 11 | 11 |
| D2S1338 | 12 | +28 | 40 | D13S317 | 8 | 8 |
| D21S11 | 19 | +27 | 46 | D7S820 | 7 | 7 |
| D8S1179 | 10 | +12 | 22 | D16S539 | 7 | 7 |
| D3S1358 | 8 | +11 | 19 | TPOX | 7 | 7 |
| vWA | 8 | +11 | 19 | TH01 | 6 | 6 |

Internal Sequence Variation in D12S391 Allele 21

One (1) size observed



Capillary electrophoresis (CE)
sizing performed with an internal
size standard

Nine (9) unique sequences observed

In 183 NIST samples

[CE 21] = AGAT[11]AGAC[10]

[CE 21] = AGAT[11]AGAC[9]AGAT[1]

[CE 21] = AGAT[12]AGAC[8]AGAT[1]

[CE 21] = AGAT[12]AGAC[9]

[CE 21] = AGAT[13]AGAC[4]AGGC
[AGAC]2AGAT[1]

[CE 21] = AGAT[13]AGAC[7]AGAT[1]

[CE 21] = AGAT[13]AGAC[8]

[CE 21] = AGAT[14]AGAC[6]AGAT[1]

[CE 21] = AGAT[14]AGAC[7]



Latest Rules and Considerations for STR Allele Nomenclature



International Society for Forensic Genetics

Forensic Science International: Genetics (May 2016) 22:54-63

Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements

Walther Parson^{a,b,*}, David Ballard^c, Bruce Budowle^{d,e}, John M. Butler^f,
Katherine B. Gettings^f, Peter Gill^{g,h}, Leonor Gusmão^{i,j,k}, Douglas R. Hares^l, Jodi A. Irwin^l,
Jonathan L. King^d, Peter de Knijff^m, Niels Morlingⁿ, Mechthild Prinz^o,
Peter M. Schneider^p, Christophe Van Neste^q, Sascha Willuweit^r, Christopher Phillips^s

^a Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

^b Forensic Science Program, The Pennsylvania State University, University Park, PA, USA

^c Faculty of Life Sciences, King's College, London, UK

^d Institute of Applied Genetics, Department of Molecular and Medical Genetics, University of North Texas Health Science Center, Fort Worth, TX, USA

^e Center of Excellence in Genomic Medicine Research (CEGMR), King Abdulaziz University, Jeddah, Saudi Arabia

^f National Institute of Standards and Technology, Gaithersburg, MD, USA

Proposed Full Description of an Allele 12 for D13S317

D13S317 Ref(11)-Chr13-GRCh38 82148025-82148068 [TATC]₁₁

D13S317[CE12]-Chr13-GRCh38 82148025-82148068
[TATC]₁₂ 82148001-A; 82148069-T

1. The reference genome assembly sequence description
- 2. Locus name and capillary electrophoresis allele name**
3. Chromosome and human genome assembly version
4. **STR repeat region coordinates [start-end] for reference allele**
5. Description of STR motifs
6. **Location of flanking region variants**

Summary of NGS/MPS

- **Additional markers** can be run simultaneously (≈ 10 times as many as current CE systems) with higher information content
 - May enable additional capabilities (e.g., phenotyping)
 - Privacy concerns with additional genomic information
- Involves **more sample preparation** steps and extensive data analysis
 - Expensive per run although cost per marker is lower
 - STR allele nomenclature challenges to keep backwards compatibility
 - Data handling and storage issues
- Primarily **still in the realm of research**
 - NIST and others are characterizing STR allele sequence variation
 - Potential advantages for mixture interpretation not demonstrated yet

Critical Challenges Faced Today

- **Success of DNA testing** → significant growth in sample submissions → sample backlogs
 - Laboratory automation and expert system data review
 - Restrictive case acceptance policies to avoid law enforcement investigator ‘swab-athons’ at crime scenes
- **Greater detection sensitivity** → more complex DNA mixtures and low-template DNA with ‘touch’ evidence
 - Probabilistic genotyping to cope with increase in data interpretation uncertainty
 - Use of a complexity threshold to avoid “skating on thin ice”

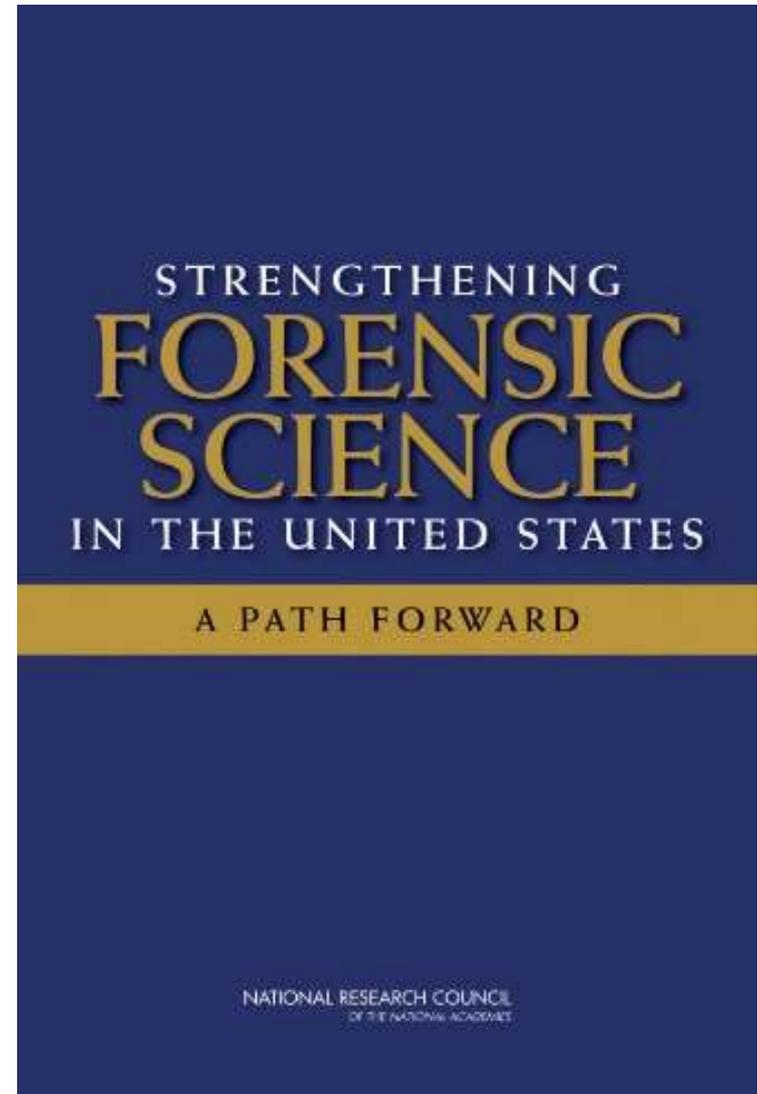
Landmark Report Gives DNA Testing a Pass

Released February 18, 2009

The U.S. National Research Council of the National Academies issued a major report on forensic science in Feb. 2009.

“With the exception of nuclear DNA analysis, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.” (p. 41)

p. 100 mentions limitations with DNA mixtures



PCAST Report Comments on Forensic DNA

Released September 20, 2016

- Supports appropriate use of single-source and simple mixture DNA analysis
- **Expresses reservations with complex DNA mixtures** (≥ 3 contributors)

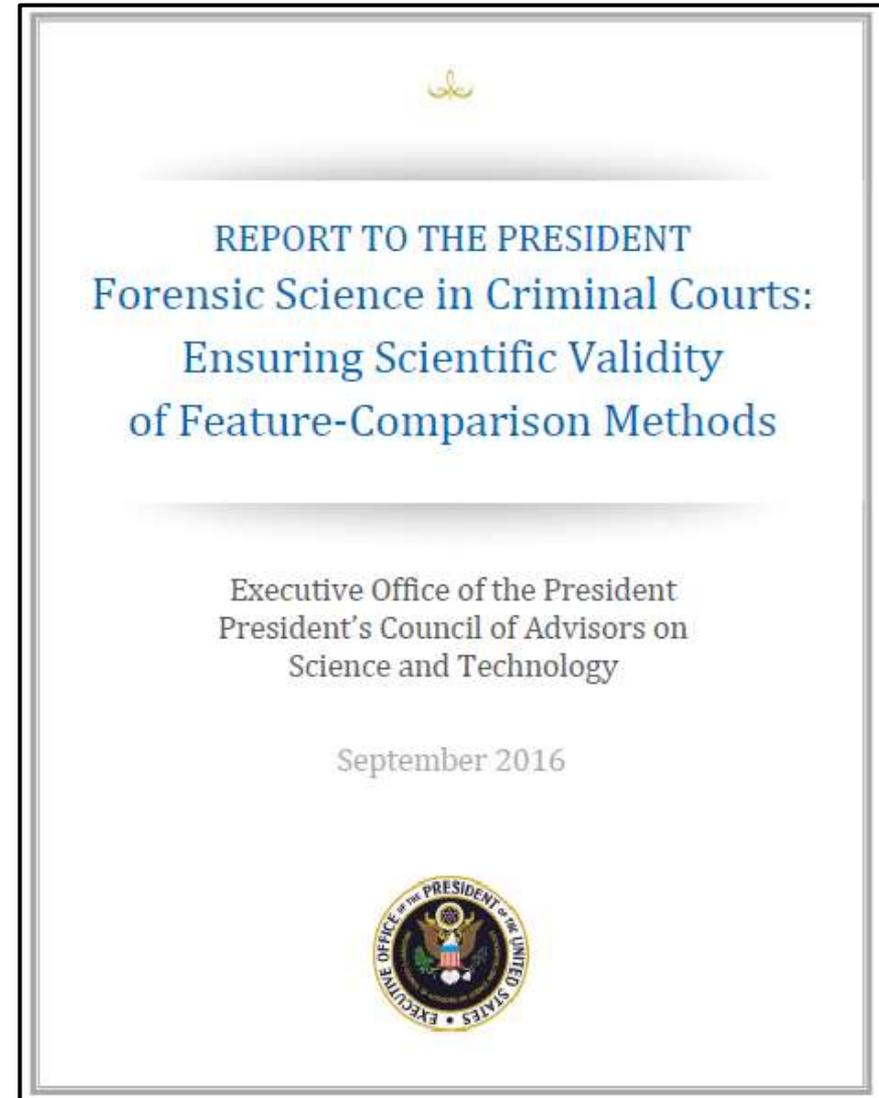
PCAST Co-Chairs



Eric Lander



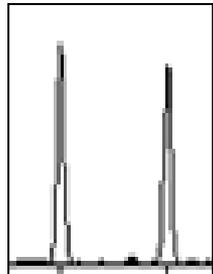
John Holdren



Math Analogy to DNA Evidence

$$2 + 2 = 4$$

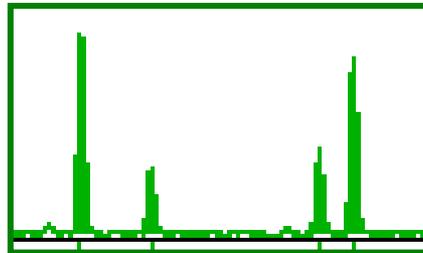
Basic Arithmetic



**Single-Source
DNA Profile**
(DNA databasing)

$$2x^2 + x = 10$$

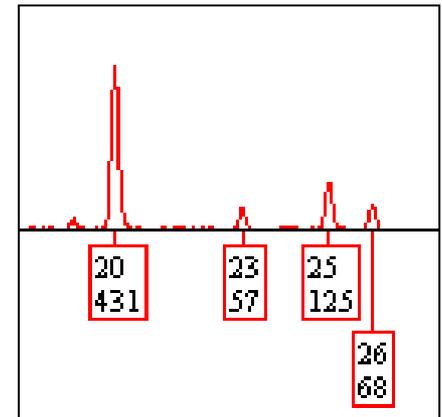
Algebra



Sexual Assault Evidence
(2-person mixture with
high-levels of DNA)

$$\int_{x=0}^{\infty} f(x) dx$$

Calculus



Touch Evidence
(>2-person, low-level,
complex mixtures
perhaps involving
relatives)

Probabilistic Genotyping

- Complex DNA mixtures with 3 or more contributors often involve low level DNA **where STR allele dropout may occur**; allele stacking and stutter artifacts also complicate interpretation
 - Currently “inconclusive” may be the only option available to analysts
- Probabilistic genotyping **uses computer simulations** to infer the likelihood of possible genotype combinations for mixture contributors
- **Several possible choices** for probabilistic genotyping software (e.g., STRmix and TrueAllele) with commercial interests at stake

The Future of Forensic DNA

is Similar to the Olympic Motto of
“Faster, Higher, Stronger”



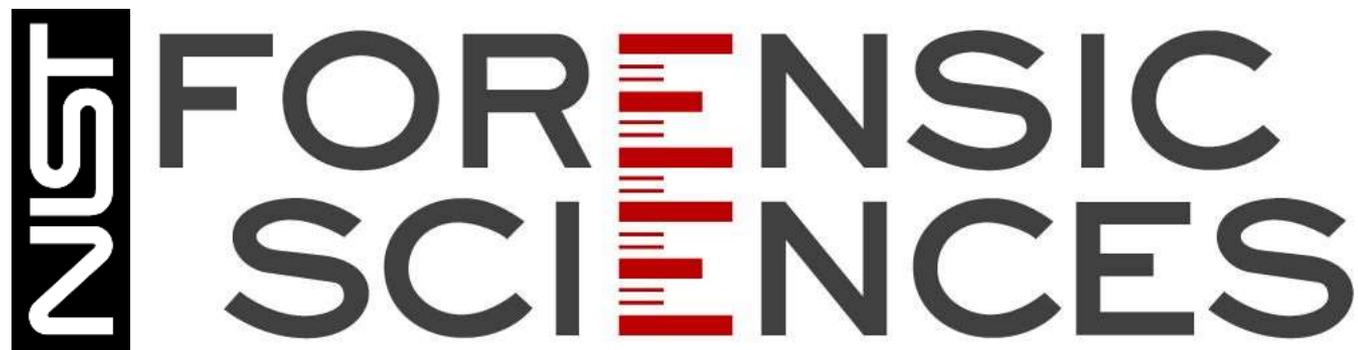
Resources

Training

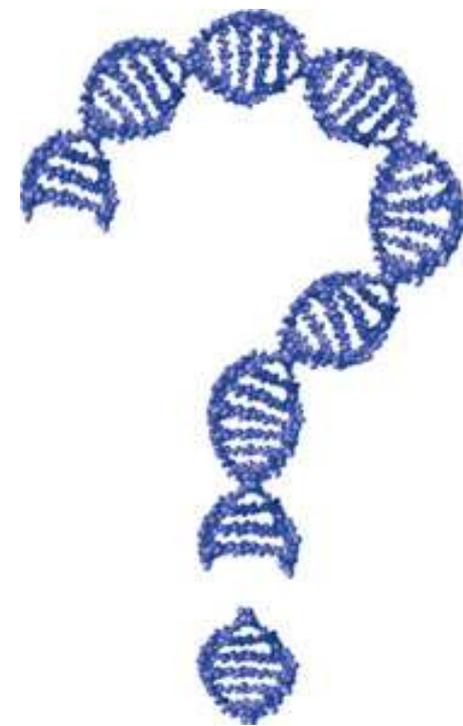
Action

National Commission on Forensic Science (NCFS):
www.justice.gov/ncfs

Organization of Scientific Area Committees (OSAC):
www.nist.gov/forensics/osac/index.cfm



www.nist.gov/forensics



+1-301-975-4049

john.butler@nist.gov